

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Nobuyuki HIRATSUKA, et al.

Application No.:

Group Art Unit:

Filed: February 4, 2004

Examiner:

For: SEARCH METHOD AND APPARATUS

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN
APPLICATION IN ACCORDANCE
WITH THE REQUIREMENTS OF 37 C.F.R. § 1.55**

Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Sir:

In accordance with the provisions of 37 C.F.R. § 1.55, the applicant(s) submit(s) herewith a certified copy of the following foreign application:

Japanese Patent Application No(s). 2003-073484

Filed: March 18, 2003

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing date(s) as evidenced by the certified papers attached hereto, in accordance with the requirements of 35 U.S.C. § 119.

Respectfully submitted,

STAAS & HALSEY LLP

Date: February 4, 2004

By: 

H. J. Staas
Registration No. 22,010

1201 New York Ave, N.W., Suite 700
Washington, D.C. 20005
Telephone: (202) 434-1500
Facsimile: (202) 434-1501

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 3 月 1 8 日
Date of Application:

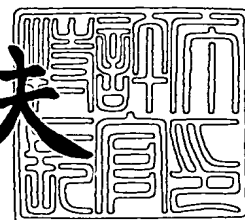
出 願 番 号 特 願 2 0 0 3 - 0 7 3 4 8 4
Application Number:
[ST. 10/C]: [J P 2 0 0 3 - 0 7 3 4 8 4]

出 願 人 富 士 通 株 式 会 社
Applicant(s):

2 0 0 3 年 1 1 月 2 5 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 3 - 3 0 9 7 2 7 9

【書類名】 特許願

【整理番号】 0350363

【提出日】 平成15年 3月18日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 19/00

【発明の名称】 検索方法及びプログラム

【請求項の数】 5

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号 富士通株式会社内

 【氏名】 平塚 信行

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号 富士通株式会社内

 【氏名】 八田 裕之

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号 富士通株式会社内

 【氏名】 渡部 勇

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号 富士通株式会社内

 【氏名】 田中 一成

【特許出願人】

 【識別番号】 000005223

 【氏名又は名称】 富士通株式会社

【代理人】**【識別番号】** 100103528**【弁理士】****【氏名又は名称】** 原田 一男**【電話番号】** 045-290-2761**【手数料の表示】****【予納台帳番号】** 076762**【納付金額】** 21,000円**【提出物件の目録】****【物件名】** 明細書 1**【物件名】** 図面 1**【物件名】** 要約書 1**【包括委任状番号】** 9909129**【プルーフの要否】** 要

【書類名】 明細書

【発明の名称】 検索方法及びプログラム

【特許請求の範囲】

【請求項 1】

ユーザによる検索条件の入力データから当該検索条件に含まれる検索語句を特定し、記憶装置に格納する語句特定ステップと、

前記検索語句及び前記検索語句の同義語の各々について、出現頻度に基づくスコアと前記検索語句又は前記検索語句の同義語を含む検索対象文書の件数とのうち少なくともいずれかである評価データを取得し、記憶装置に格納する評価データ取得ステップと、

前記検索語句及び前記検索語句の同義語と対応する前記評価データとを、1又は複数の前記検索語句及び前記検索語句の同義語を選択可能な態様で前記ユーザに提示する提示ステップと、

前記ユーザにより選択された前記検索語句又は前記検索語句の同義語を含む検索対象文書に関するデータを、前記ユーザに提示する結果提示ステップと、

を含むコンピュータにより実行される検索方法。

【請求項 2】

前記評価データ取得ステップが、

前記検索語句から同義語を抽出するステップと、

前記検索語句及び前記検索語句の同義語を用いて検索対象文書群を検索することにより、前記検索語句又は前記検索語句の同義語を含む検索対象文書の件数と前記検索語句及び前記検索語句の各々の同義語の第1の出現回数とのうち少なくともいずれかを計数するステップと、

を含む請求項 1 記載の検索方法。

【請求項 3】

前記評価データ取得ステップが、

前記検索条件として入力された文章における前記検索語句の第2の出現回数を計数するステップと、

前記検索語句の第2の出現回数と前記検索語句及び前記検索語句の同義語の各

々の第1の出現回数とを用いて、前記出現頻度に基づくスコアを計算するステップと、

をさらに含む請求項2記載の検索方法。

【請求項4】

前記提示ステップが、

前記検索語句及び前記検索語句の同義語の評価データが所定の条件を満たすか判断するステップと、

前記評価データが所定の条件を満たす前記検索語句又は前記検索語句の同義語については予め選択された状態で、前記評価データが所定の条件を満たさない前記検索語句又は前記検索語句の同義語については未選択の状態で前記ユーザに提示するステップと、

を含む請求項1乃至3のいずれか1つ記載の検索方法。

【請求項5】

ユーザによる検索条件の入力データから当該検索条件に含まれる検索語句を特定し、記憶装置に格納する語句特定ステップと、

前記検索語句及び前記検索語句の同義語の各々について、出現頻度に基づくスコアと前記検索語句又は前記検索語句の同義語を含む検索対象文書の件数とのうち少なくともいずれかである評価データを取得し、記憶装置に格納する評価データ取得ステップと、

前記検索語句及び前記検索語句の同義語と対応する評価データとを、1又は複数の前記検索語句及び前記検索語句の同義語を選択可能な態様で前記ユーザに提示する提示ステップと、

前記ユーザにより選択された前記検索語句又は前記検索語句の同義語を含む検索対象文書に関するデータを、前記ユーザに提示する結果提示ステップと、

をコンピュータに実行させるプログラム。

【発明の詳細な説明】

【0001】

【発明が属する技術分野】

本発明は、文書データの検索技術に関する。

【0002】**【従来の技術】**

従来の検索システムでは、検索したいテーマに関する検索タームを指定して検索するのが一般的であった。例えば、特許情報の検索システムでは「キーワード」や「IPC」、「出願人」などの様々な検索タームを駆使して検索するのが一般的である。しかし、このような検索手法では、効果的な検索タームを思いつくこと自体がノウハウであり、ある程度の熟練者でないと効果的な検索ができないという問題があった。

【0003】

そこで、上述のような問題を解決するために、近年の検索システムでは、利用者が入力した文章から、その入力文に類似するものを検索し、類似度順に並べて表示する検索手法（以下「概念検索」と呼ぶ）を利用して、初心者でも簡単に目的の文献を探し出すことができるようになってきている。

【0004】

この概念検索では、利用者が入力した文章から、形態素解析により語句を抽出し、入力文から抽出された各語句を用いて、データベースに管理されている各文献における抽出語句群の出現頻度と、データベース全体での抽出語句群の出現頻度とを利用して、例えばTF/IDF法などにより抽出語句の重みを計算し、重みに従って順番に並べて表示する。

【0005】

また、特開平9-297766号公報には、以下のような類似文書検索装置が開示されている。すなわち、形態素解析部により認識された入力文書中のキーワードの個数を計数するキーワードカウント部、文書に含まれるキーワードを意味分類毎に仕訳するキーワード意味分類決定部、意味分類に応じた重要度と各意味分類に属するキーワードの個数に依存する評価値を付与する意味分類評価値決定部、及び評価値に基づいて各参照用文書毎に類似度を付与する文書類似度決定部を含む。

【0006】**【特許文献1】**

特開平 9-297766 号公報

【0007】

【発明が解決しようとする課題】

このように概念検索を利用することにより、初心者でも比較的簡単に類似する文献を検索できるようになったが、概念検索で一定以上の検索精度を達成するためには入力する文章の精度、すなわち類似度の計算に利用する語句（抽出語句）の精度が重要となってくる。従って、同義語、異表記など同じ意味で表現が異なる語句（以下同義語と呼ぶ）の考慮がない場合には検索精度が落ちてしまう。例えば高速道路のみ抽出された場合にはハイウェイが落ちてしまっていると検索精度が落ちる。また、検索テーマに直接的に影響しない語句があることで結果が散漫になってしまう場合もある。さらに、影響が強すぎる語句が含まれることで結果が偏ってしまう場合もある。

【0008】

また特開平 9-297766 号公報のように意味分類に属するキーワードの個数に依存する評価値を計算する方法もあるが、この評価方法では意味分類毎に重要度を設定して評価値を計算することになるため、意味分類が適切であること及び意味分類毎の重要度が適切に設定されていることが前提となる。しかし、いずれの場合においてもそれらの設定が適切であるということとはありえない。

【0009】

従って、本発明の目的は、よりの確な検索結果を得るためにユーザを適切にガイドする検索処理技術を提供することである。

【0010】

【課題を解決するための手段】

本発明に係る検索方法は、ユーザによる検索条件の入力データから当該検索条件に含まれる検索語句を特定し、記憶装置に格納する語句特定ステップと、検索語句及び当該検索語句の同義語の各々について、出現頻度に基づくスコアと検索語句又は当該検索語句の同義語を含む検索対象文書の件数とのうち少なくともいずれかである評価データを取得し、記憶装置に格納する評価データ取得ステップと、検索語句及び当該検索語句の同義語と対応する評価データとを、1又は複数

の検索語句及び当該検索語句の同義語を選択可能な態様でユーザに提示する提示ステップと、ユーザにより選択された検索語句又は当該検索語句の同義語を含む検索対象文書に関するデータを、ユーザに提示する結果提示ステップとを含む。

【0011】

このような検索方法を用いることにより、単に検索条件に含まれる検索語句だけではなく同義語を含めて検索でき、さらに検索対象文書との関連性を表す評価データを提示して語句の選択についてユーザをガイドするため、ユーザにとって適切な検索が行われるようになる。

【0012】

なお、上で述べた評価データ取得ステップが、検索語句から同義語を抽出するステップと、検索語句及び当該検索語句の同義語を用いて検索対象文書群を検索することにより、検索語句又は当該検索語句の同義語を含む検索対象文書の件数と検索語句及び当該検索語句の同義語の各々の第1の出現回数とのうち少なくともいずれかを計数するステップとを含むようにしてもよい。別途各語句について予め検索及び計数を行っておき、当該計数結果を用いるようにしても良い。

【0013】

さらに、上で述べた評価データ取得ステップが、検索条件として入力された文章における検索語句の第2の出現回数を計数するステップと、検索語句の第2の出現回数と検索語句及び当該検索語句の同義語の各々の第1の出現回数とを用いて、出現頻度に基づくスコアを計算するステップとをさらに含むようにしてもよい。このように第1及び第2の出現回数を用いることにより、語句の重要性を入力文章と検索対象文書群との相対的な関係から導き出すことができ、ユーザはより語句の選択を的確に行いやすくなる。

【0014】

なお、上述の方法はプログラム及びコンピュータにて実施することができ、当該プログラムは、例えばフレキシブルディスク、CD-ROM、光磁気ディスク、半導体メモリ、ハードディスク等の記憶媒体又は記憶装置に格納される。また、ネットワークなどを介してデジタル信号として配信される場合もある。尚、中間的な処理結果はメモリに一時保管される。

【0015】**【発明の実施の形態】**

図1に本発明のシステム概要図を示す。例えばインターネットやLAN (Local Area Network) であるネットワーク1には、例えばパーソナルコンピュータでウェブ (Web) ブラウザ機能を有するユーザ端末3及び7と、Webサーバ機能を有しており本実施の形態における主たる処理を実施する検索サーバ5とが接続されている。検索サーバ5は、検索条件処理部51と、検索処理部52と、検索後処理部53とを含み、ファイル格納部54と文献群データベース (DB) 55を管理する。

【0016】

図1に示したシステムの処理内容を図2乃至図11を用いて説明する。検索者は、ユーザ端末3を操作して、検索条件入力ページへアクセスさせる (ステップS1)。検索サーバ5の検索条件処理部51は、ユーザ端末3からのアクセスに応じて、検索条件入力ページのデータをユーザ端末3に送信する (ステップS3)。ユーザ端末3は、検索サーバ5から検索条件入力ページ・データを受信し、表示装置に表示する (ステップS5)。例えば図3のような画面が表示される。

【0017】

図3は、特許検索の例を示しており、全公報、公開公報、登録公報等の検索対象を選択するための検索対象選択欄301と、入力文章からの同義語を展開した場合に当該展開語句を検索者が選択するか選択しないかを選択入力する選択欄302と、検索ボタン303と、条件式をクリアするための条件式クリアボタン304と、検索用の文章入力欄305と、他の検索項目指定欄306及び309と、他の検索項目の検索キーワード入力欄307及び310と、検索キーワードについての関係 (すべてを含む、いずれかを含むなど) を指定するための選択欄308及び311と、公報発行期間の指定欄312と、検索結果の処理対象選択欄313と、表示件数の選択欄314と、処理結果表示欄315とが含まれる。

【0018】

ユーザは、図3のような画面を見て、検索対象を選択し、文章 (図3では「高速道路で停止することなく料金を支払う方法」) を入力し、他の検索項目及び検

索キーワードの関係を選択し且つ検索キーワードを入力し、公報発行日を入力し、そして検索ボタン 303 をクリックする。必要な部分のみ入力するようにしても良い。ユーザ端末 3 は、検索者による例えば入力文章を含む検索条件の入力を受け付け、検索サーバ 5 に送信する（ステップ S 7）。検索サーバ 5 の検索条件処理部 51 は、ユーザ端末 3 から例えば入力文章を含む検索条件を受信し、一旦ワークメモリ領域（例えばメインメモリなどに確保された領域）に格納する（ステップ S 9）。検索条件処理部 51 は、入力文章に対して周知の形態素解析を行って語句を抽出し、抽出語句ファイルに登録する（ステップ S 11）。上で述べた文章が入力された場合には図 4 に示すように「高速道路」「停止」「料金」「支払」「方法」という語句（抽出語句）が抽出され、抽出語句ファイルに登録される。

【0019】

そして検索条件処理部 51 及び検索処理部 52 は、抽出語句の文献数及びスコア取得処理を実施する（ステップ S 13）。この処理について図 5 を用いて詳細に説明する。検索条件処理部 51 は、1つの抽出語句を抽出語句ファイルからワークメモリ領域に読み出す（ステップ S 41）。そして、検索処理部 52 は、文献群 DB 55 を当該抽出語句で検索し、当該抽出語句について該当文献数及び出現頻度を計数し、一旦ワークメモリ領域に格納する（ステップ S 43）。なお、各語句で文献群 DB 55 を予め検索して該当文献数及び出現頻度を計数しておき、当該計数結果をこの段階で読み出すようにしても良い。また、入力文章を抽出語句で検索し、出現頻度を計数し、一旦ワークメモリ領域に格納する（ステップ S 44）。そして、検索条件処理部 51 は、抽出語句のスコアを計算し、ワークメモリ領域に格納する（ステップ S 45）。本実施の形態における抽出語句のスコアは、 $\{ (入力文章における抽出語句の出現頻度) / (文献群 DB 55 における抽出語句の出現頻度) \}$ で計算される。検索条件処理部 51 は、このように計数又は計算された該当文献数及びスコアを、抽出語句に対応して第 2 抽出語句ファイルに書き込む（ステップ S 47）。

【0020】

第 2 抽出語句ファイルの一例を図 6 に示す。図 6 のファイル構成例では、語句

の列 3 2 1 と、ヒット文献数（該当文献数）の列 3 2 2 と、スコアの列 3 2 3 と、選択フラグの列 3 2 4 とが含まれている。ステップ S 4 7 では、語句の列 3 2 1 と、ヒット文献数の列 3 2 2 と、スコアの列 3 2 3 とに値を登録する。

【0021】

そして、検索条件処理部 5 1 は、同義語ファイルを参照して、抽出語句の同義語を抽出する（ステップ S 4 9）。同義語ファイルには、例えば図 7 に示すように元の語句の列 3 4 1 と、同義語の列 3 4 2 とが設けられており、特定の語句（元の語句）に対応して 1 又は複数の同義語が登録されている。従って、元の語句の列 3 4 1 を抽出語句で検索し、同義語の列 3 4 2 の対応する語句を読み出す。

【0022】

検索処理部 5 2 は、文献群 DB 5 5 を 1 つの同義語で検索し、当該同義語について該当文献数及び出現頻度を計数する（ステップ S 5 1）。なお、各語句で文献群 DB 5 5 を予め検索して該当文献数及び出現頻度を計数しておき、当該計数結果をこの段階で読み出すようにしても良い。そして、検索条件処理部 5 1 は、同義語のスコアを計算し、ワークメモリ領域に格納する（ステップ S 5 3）。本実施の形態における同義語のスコアは、 $\{（同義語に対応する抽出語句（元の語句）の、入力文章における出現頻度）／（文献群 DB 5 5 における抽出語句の出現頻度）\}$ で計算される。検索条件処理部 5 1 は、このように計数又は計算された該当文献数及びスコアを、同義語に対応して第 2 抽出語句ファイル（図 6）に書き込む（ステップ S 5 5）。ステップ S 5 5 では、語句の列 3 2 1 と、ヒット文献数の列 3 2 2 と、スコアの列 3 2 3 とに値を登録する。

【0023】

そしてステップ S 4 1 において特定された抽出語句に対応する全ての同義語について処理したか判断する（ステップ S 5 7）。もし、未処理の同義語が存在する場合にはステップ S 4 9 に戻る。一方、全ての同義語についての処理が終了した場合にはステップ S 5 9 に移行する。そして未処理の抽出語句が存在するか判断する（ステップ S 5 9）。未処理の抽出語句が存在する場合には、ステップ S 4 1 に戻る。全ての抽出語句について処理が終了すれば元の処理に戻る。

【0024】

図2の説明に戻って、検索条件処理部51は、閾値チェック処理を実施する（ステップS15）。この閾値チェック処理について図8を用いて説明する。検索条件処理部51は、閾値ファイルから閾値を読み出す（ステップS61）。閾値ファイルの一例を図9に示す。図9のファイル構成例では、項目の列351と閾値の列352とが設けられており、文献数についての閾値（例えば1000）とスコアについての閾値（0.300）とが登録されている。そして、第2抽出語句ファイルから1つの語句のデータを読み出す（ステップS63）。この語句の該当文献数が文献数についての閾値を超えているか判断する（ステップS65）。該当文献数が多いと検索結果が散漫になってしまうため、この段階でチェックする。この語句の該当文献数が文献数についての閾値以下である場合には、第2抽出語句ファイルに選択フラグをセットする（ステップS69）。図6に示した例では、選択フラグの列324の対応するフラグをONにセットする。なお、デフォルトをOFFにしておく。そしてステップS71に移行する。

【0025】

一方、この語句の該当文献数が文献数についての閾値を超えている場合には、この語句のスコアがスコアについての閾値を超えているか判断する（ステップS67）。スコアが低いのは、文献群DB55における当該語句の出現頻度が高い場合又は入力文章において出現頻度が低い場合若しくはその両方である。一方、スコアが高いのは、文献群DB55における当該語句の出現頻度が低い場合又は入力文章において出現頻度が高い場合若しくはその両方である。このようにスコアによって、当該語句がこの検索において特徴的なものか否か、若しくはこの検索における当該語句の重要性が高いか否かを判断することができる。本実施の形態では、固定的な重要度や重み付けではなく、入力文章と文献群DB55との相対的な関係から語句の重要性等が導き出されるので、より状況にあった数値をユーザに提示できるようになる。

【0026】

この語句のスコアがスコアについての閾値を超えている場合にはステップS69に移行する。一方、この語句のスコアがスコアについての閾値以下である場合には、未処理の語句が存在するか判断する（ステップS71）。未処理の語句が

存在する場合にはステップ S 6 3 に戻る。一方、全ての語句について処理が完了している場合には元の処理に戻る。

【0027】

このようにして、検索サーバ 5 は、検索者に対して検索に用いることを推奨する語句を自動的に選定する。従って、検索者は、初心者であっても、的確な語句を選ぶことができるようになる。

【0028】

図 2 の処理に戻って、検索条件処理部 5 1 は、第 2 抽出語句ファイル (図 6) のデータを用いて、抽出語句及び同義語と対応するスコア及び文献数のデータを含む抽出語句選択ページのデータを生成し、ユーザ端末 3 に送信する (ステップ S 1 7)。ユーザ端末 3 は、検索サーバ 5 から抽出語句選択ページのデータを受信し、表示装置に表示する (ステップ S 1 9)。例えば図 1 0 に示すような画面が表示される。

【0029】

図 1 0 の例では、検索ボタン 3 6 1 と、チェックボックスの列 3 6 2 と、抽出語句 (同義語を含む) の列 3 6 3 と、スコアの列 3 6 4 と、文献数の列 3 6 5 とが設けられている。なお、第 2 抽出語句ファイルの選択フラグの列 3 2 4 においてフラグがセットされている語句については、デフォルトでチェックボックスにチェックが付されている。検索者は、このチェックをはずすことも可能であるし、さらにチェックを付すことも可能である。このように本実施の形態では、スコア及び文献数にて検索者が的確な語句を選択して的確な検索を行えるようにガイドしている。

【0030】

検索者は、スコアの値や文献数を参照して、チェックを付すべき語句及びチェックをはずす語句を選択する。そして、チェックボックスにチェックを付したり、チェックをはずしたりした後に、検索ボタン 3 6 1 をクリックする。ユーザ端末 3 は、検索者から語句選択入力 (選択をはずす入力を含む) を受け付け (ステップ S 2 1)、ユーザ端末 3 は、選択された語句についてのデータを検索サーバ 5 に送信する (ステップ S 2 3)。検索サーバ 5 の検索処理部 5 2 は、ユーザ端

末 3 から選択された語句についてのデータを受信し、一旦ワークメモリ領域に格納する（ステップ S 2 5）。そして、選択された語句を用いて文献群 DB 5 5 を検索する（ステップ S 2 7）。なお、上で行った検索の結果を保持しておき、この段階にて当該結果を読み出すようにしても良い。さらに、各語句について行われた検索結果を保持しておき、それを読み出すようにしても良い。そして、検索後処理部 5 3 は、検索結果である各文献についてスコアを計算し、ランク付けを行い、例えばワークメモリ領域に格納する（ステップ S 2 9）。本実施の形態では、文献についてのスコアは、 $\{ (\text{文献における、検索者により選択された語句の出現頻度}) / (\text{文献群 DB 5 5 における、検索者により選択された語句の出現頻度}) \}$ の総和にて計算される。このスコアの値の大きい順にランク付けがなされる。

【0031】

検索後処理部 5 3 は、ランク付け結果を用いて検索結果ページ・データを生成し、ユーザ端末 3 に送信する（ステップ S 3 1）。ユーザ端末 3 は、検索サーバ 5 から検索結果ページ・データを受信し、表示装置に表示する（ステップ S 3 3）。例えば図 1 1 に示すような画面が表示される。

【0032】

図 1 1 の例では、図 3 に示した画面の処理結果表示欄 3 1 5 に処理結果 3 7 1 が表示されている。処理結果 3 7 1 は、文献の選択を示すためのチェックボックスの列 3 7 2 と、ランキングの列 3 7 3 と、文献番号及び文献内容の列 3 7 4 とが設けられている。このようにより入力文章と関連性が高いとされる文献順に検索結果が提示されるため、ユーザはより文献の特定がしやすくなる。

【0033】

以上本発明の一実施の形態を説明したが、本発明はこれに限定されるものではない。例えば、図 1 に示した機能ブロックは必ずしもプログラムモジュールに対応するものではない。また、図 1 ではクライアント・サーバ環境での実施の形態を説明したが、検索サーバ 5 の機能並びに文献群 DB 5 5 並びにファイル格納部 5 7 を備えた端末を構成することも可能である。

【0034】

またスコアの計算方法についても一例であって、他の方法にて計算するようにしても良い。図3、図10及び図11の画面構成は一例であって、他の画面構成を採用することも可能である。処理結果については別ウインドウにて示すようにしても良い。さらに、スコアと文献数を両方ともユーザに提示する例を示したが、いずれか一方のみをユーザに提示することも可能である。

【0035】

(付記1)

ユーザによる検索条件の入力データから当該検索条件に含まれる検索語句を特定し、記憶装置に格納する語句特定ステップと、

前記検索語句及び前記検索語句の同義語の各々について、出現頻度に基づくスコアと前記検索語句又は前記検索語句の同義語を含む検索対象文書の件数とのうち少なくともいずれかである評価データを取得し、記憶装置に格納する評価データ取得ステップと、

前記検索語句及び前記検索語句の同義語と対応する前記評価データとを、1又は複数の前記検索語句及び前記検索語句の同義語を選択可能な態様で前記ユーザに提示する提示ステップと、

前記ユーザにより選択された前記検索語句又は前記検索語句の同義語を含む検索対象文書に関するデータを、前記ユーザに提示する結果提示ステップと、

を含むコンピュータにより実行される検索方法。

【0036】

(付記2)

前記語句特定ステップが、

前記検索条件として入力された文章から形態素解析により検索語句を抽出するステップ

を含む付記1記載の検索方法。

【0037】

(付記3)

前記評価データ取得ステップが、

前記検索語句から同義語を抽出するステップと、

前記検索語句及び前記検索語句の同義語を用いて検索対象文書群を検索することにより、前記検索語句又は前記検索語句の同義語を含む検索対象文書の件数と前記検索語句及び前記検索語句の同義語の各々の第 1 の出現回数とのうち少なくともいずれかを計数するステップと、

を含む付記 1 又は 2 記載の検索方法。

【 0 0 3 8 】

(付記 4)

前記評価データ取得ステップが、

前記検索条件として入力された文章における前記検索語句の第 2 の出現回数を計数するステップと、

前記検索語句の第 2 の出現回数と前記検索語句及び前記検索語句の同義語の各々の第 1 の出現回数とを用いて、前記出現頻度に基づくスコアを計算するステップと、

をさらに含む付記 3 記載の検索方法。

【 0 0 3 9 】

(付記 5)

前記提示ステップが、

前記検索語句及び前記検索語句の同義語の評価データが所定の条件を満たすか判断するステップと、

前記評価データが所定の条件を満たす前記検索語句又は前記検索語句の同義語については予め選択された状態で、前記評価データが所定の条件を満たさない前記検索語句又は前記検索語句の同義語については未選択の状態で前記ユーザに提示するステップと、

を含む付記 1 乃至 4 のいずれか 1 つ記載の検索方法。

【 0 0 4 0 】

(付記 6)

前記所定の条件が、

前記検索語句又は前記検索語句の同義語を含む検索対象文書の件数が第 1 の閾値未満、又は前記検索語句又は前記検索語句の同義語の前記出現頻度に基づくス

コアが第2の閾値以上である

ことを特徴とする付記1記載の検索方法。

【0041】

(付記7)

前記結果提示ステップが、

前記ユーザにより選択された前記検索語句又は前記検索語句の同義語を含む検索対象文書における、前記ユーザにより選択された前記検索語句又は前記検索語句の同義語の第3の出現回数を計数するステップと、

前記第3の出現回数を用いて計算される数値の順番にて前記検索対象文書を提示するステップと、

を含む付記1記載の検索方法。

【0042】

(付記8)

ユーザによる検索条件の入力データから当該検索条件に含まれる検索語句を特定し、記憶装置に格納する語句特定ステップと、

前記検索語句及び前記検索語句の同義語の各々について、出現頻度に基づくスコアと前記検索語句又は前記検索語句の同義語を含む検索対象文書の件数とのうち少なくともいずれかである評価データを取得し、記憶装置に格納する評価データ取得ステップと、

前記検索語句及び前記検索語句の同義語と対応する評価データとを、1又は複数の前記検索語句及び前記検索語句の同義語を選択可能な態様で前記ユーザに提示する提示ステップと、

前記ユーザにより選択された前記検索語句又は前記検索語句の同義語を含む検索対象文書に関するデータを、前記ユーザに提示する結果提示ステップと、

をコンピュータに実行させるプログラム。

【0043】

(付記9)

ユーザによる検索条件の入力データから当該検索条件に含まれる検索語句を特定し、記憶装置に格納する手段と、

前記検索語句及び前記検索語句の同義語の各々について、出現頻度に基づくスコアと前記検索語句又は前記検索語句の同義語を含む検索対象文書の件数とのうち少なくともいずれかである評価データを取得し、記憶装置に格納する手段と、

前記検索語句及び前記検索語句の同義語と対応する評価データとを、1又は複数の前記検索語句及び前記検索語句の同義語を選択可能な態様で前記ユーザに提示する手段と、

前記ユーザにより選択された前記検索語句又は前記検索語句の同義語を含む検索対象文書に関するデータを、前記ユーザに提示する手段と、

を有する検索装置。

【 0 0 4 4 】

【発明の効果】

以上述べたように本発明によれば、よりの確な検索結果を得るためにユーザを適切にガイドすることができる。

【図面の簡単な説明】

【図 1】

本発明の一実施の形態における機能ブロックを示す図である。

【図 2】

本発明の実地の形態におけるメインの処理フローを示す図である。

【図 3】

検索条件入力画面の一例を示す図である。

【図 4】

抽出語句ファイルに格納されるデータの一例を示す図である。

【図 5】

抽出語句の文献数及びスコア取得処理の処理フローを示す図である。

【図 6】

第 2 抽出語句ファイルに格納されるデータの一例を示す図である。

【図 7】

同義語ファイルに格納されるデータの一例を示す図である。

【図 8】

閾値チェック処理の処理リフローを示す図である。

【図 9】

閾値ファイルの一例を示す図である。

【図 1 0】

抽出語句選択画面の一例を示す図である。

【図 1 1】

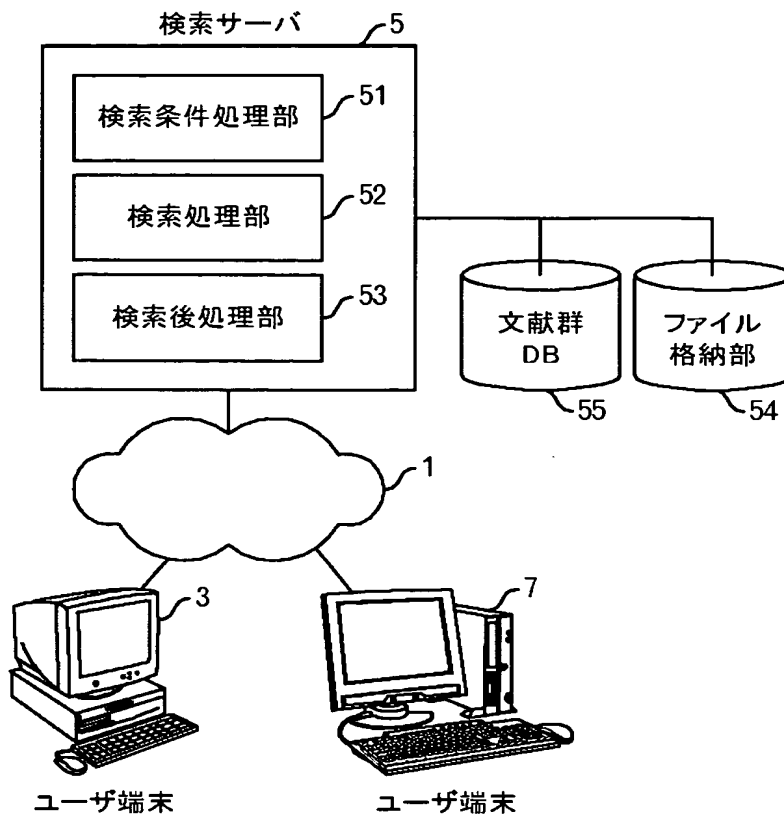
検索結果表示画面の一例を示す図である。

【符号の説明】

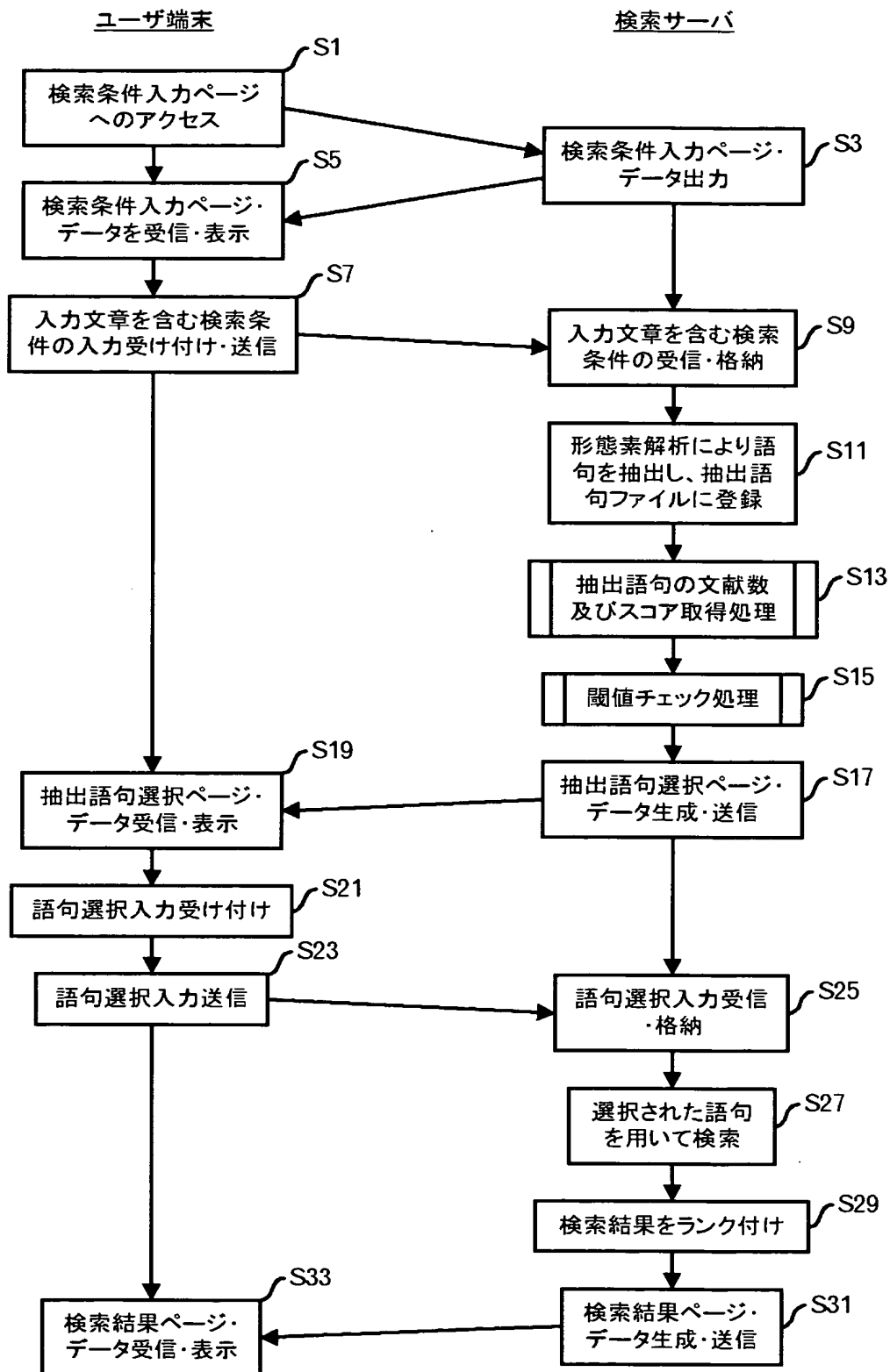
- 1 ネットワーク 3, 7 ユーザ端末
- 5 検索サーバ
- 5 1 検索条件処理部 5 2 検索処理部
- 5 3 検索後処理部 5 4 ファイル格納部
- 5 5 文献群 D B

【書類名】 図面

【図 1】



【図 2】



【図 3】

303 検索対象

301 全公開(公開/登録) ▼

304 文章からの展開語句 ▼

302 選択する ▼

305 文章

307 文章を入力して検索ボタンをクリックしてください。

308 高速道路で停止することなく料金を支払う方法

306 出願人(部分) ▼

309 出願人(部分) ▼

310 公報発行日

西暦 ▼

年

月

日

~

年

月

日

311 すべてを含む ▼

312 すべてを含む ▼

処理対象: すべて ▼

表示件数: 20件 ▼

313

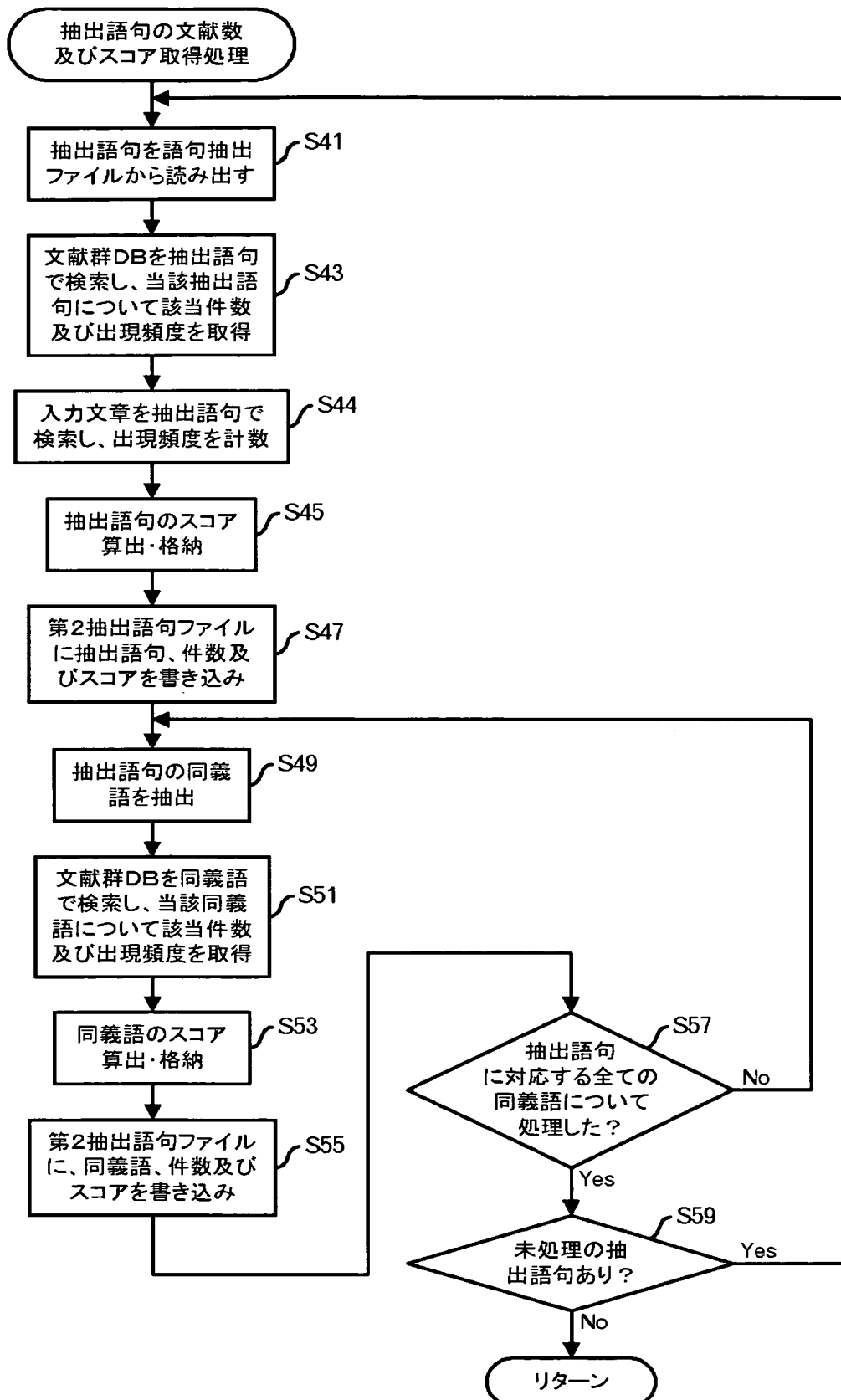
314

315

【図 4】

高速道路	停止	料金	支払	方法
------	----	----	----	----

【図 5】



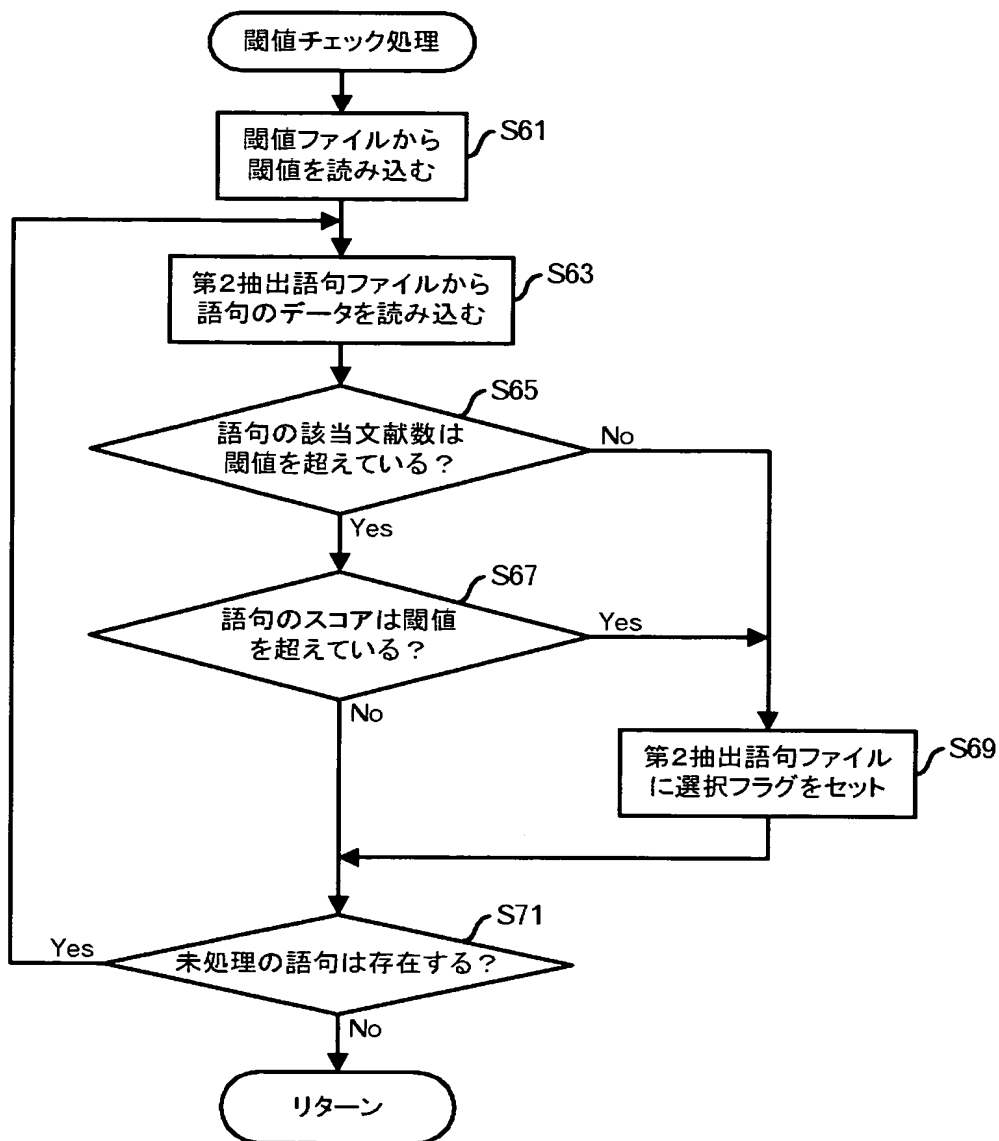
【図 6】

語句	ヒット文献数	スコア	選択フラグ
高速道路	1000	0.800	ON
高速道	800	0.700	ON
ハイウェイ	600	0.200	OFF
有料道	400	0.500	ON
道	100,000	0.500	OFF
・	・	・	・
方法	4,000	0.010	OFF

【図 7】

元語句	同義語
高速道路	高速道
	ハイウェイ
	有料道
	道
停止	ストップ
	止める
	中断
・	・
・	・
・	・

【図 8】



【図 9】

項目	閾値
文献数	1000
スコア	0.300

【図 10】

抽出語句画面				
361 検索	362	363 抽出語句	364 スコア	365 文献数
<input checked="" type="checkbox"/>	高速道路	0.800	1000件	
<input checked="" type="checkbox"/>	高速道	0.700	800件	
<input type="checkbox"/>	ハイウェイ	0.200	600件	
<input checked="" type="checkbox"/>	有料道	0.500	400件	
<input checked="" type="checkbox"/>	道	0.500	100000件	
<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>	方法	0.010	4000件	
<input type="checkbox"/>				

【図 1 1 1】

検索対象

全公報(公開／登録) ▼

文章からの展開語句

選択する ▼

検索

条件式クリア

文章を入力して検索ボタンをクリックしてください。

文章

高速道路で停止することなく料金を支払う方法

出願人(部分) ▼

すべてを含む ▼

出願人(部分) ▼

すべてを含む ▼

公報発行日

西暦 ▼

年

月

日

年

月

日

検索結果: 192,932件

処理対象: すべて ▼

表示件数: 20件 ▼

次一覧

371

☐ 00001

特開2000-014427

【名称】通行券のホルダー(57)【要約】【課題】モーターバイクで高速道路などの有料道路を走行するとき、…

☐ 00002

特開2000-067383

【名称】車載用ナビゲーション装置(57)【要約】【課題】高速道路の渋滞を示す矢印と一般道路の渋滞を…

☐ 00003

特開2000-142249

【名称】自動車用コインケース(57)【要約】【課題】コインの取り出しと収納が便利な自動車用コインケース…

☐ 00004

特開2000-090390

【名称】高速道路の走行違反警告システム(57)【要約】【課題】高速道路における走行違反車両に対して…

☐ 00005

特開2000-067380

【名称】地図表示方法(57)【要約】【課題】一見して高速道路を認識できるように、あるいは、一見してVICS…

372

373

374

【書類名】 要約書

【要約】

【課題】

よりの確な検索結果を得るためにユーザを適切にガイドする。

【解決手段】

ユーザによる検索条件の入力データから当該検索条件に含まれる検索語句を特定するステップと、検索語句及びその同義語の各々について、出現頻度に基づくスコアと当該検索語句又は同義語を含む検索対象文書の件数とのうち少なくともいずれかである評価データを取得するステップと、検索語句及びその同義語と対応する評価データとを、1又は複数の検索語句及びその同義語を選択可能な態様でユーザに提示するステップと、ユーザにより選択された検索語句又はその同義語を含む検索対象文書に関するデータを、ユーザに提示するステップとを含む。単に検索条件に含まれる検索語句だけではなく同義語を含めて検索でき、さらに検索対象文書との関連性を表す評価データを提示して語句の選択についてユーザをガイドするため、ユーザにとって適切な検索が行われる。

【選択図】 図1

特願 2 0 0 3 - 0 7 3 4 8 4

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 2 2 3]

1 . 変更年月日

1 9 9 6 年 3 月 2 6 日

[変更理由]

住所変更

住 所

神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号

氏 名

富士通株式会社